



## **ViralORFeome: an integrated database to generate a versatile collection of viral ORFs.**

J. Pellet, L. Tafforeau, M. Lucas-Hourani, V. Navratil, L. Meyniel, Guillaume Achaz, A. Guironnet-Paquet, A. Aublin-Gex, G. Caignard, P. Cassonnet, et al.

### **► To cite this version:**

J. Pellet, L. Tafforeau, M. Lucas-Hourani, V. Navratil, L. Meyniel, et al.. ViralORFeome: an integrated database to generate a versatile collection of viral ORFs.. *Nucleic Acids Research*, 2010, 38 (Issue suppl1: Database issue), pp.D371-8. 10.1093/nar/gkp1000 . pasteur-00455342

**HAL Id: pasteur-00455342**

**<https://hal-pasteur.archives-ouvertes.fr/pasteur-00455342>**

Submitted on 11 Feb 2010

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution - NonCommercial| 4.0 International License

# ViralORFeome: an integrated database to generate a versatile collection of viral ORFs

J. Pellet<sup>1,2,3</sup>, L. Tafforeau<sup>1,2,3</sup>, M. Lucas-Hourani<sup>4</sup>, V. Navratil<sup>2,3,5</sup>, L. Meyniel<sup>1,2,3</sup>, G. Achaz<sup>6</sup>, A. Guironnet-Paquet<sup>1,2,3,7</sup>, A. Aublin-Gex<sup>1,2,3</sup>, G. Caignard<sup>4</sup>, P. Cassonnet<sup>8</sup>, A. Chaboud<sup>2,3,9</sup>, T. Chantier<sup>1,2,3</sup>, A. Deloire<sup>1,2,3</sup>, C. Demeret<sup>8</sup>, M. Le Breton<sup>1,2,3</sup>, G. Neveu<sup>8</sup>, L. Jacotot<sup>6</sup>, P. Vaglio<sup>6</sup>, S. Delmotte<sup>2,3,7</sup>, C. Gautier<sup>2,3,7</sup>, C. Combet<sup>2,3,10</sup>, G. Deleage<sup>2,3,10</sup>, M. Favre<sup>8</sup>, F. Tangy<sup>4</sup>, Y. Jacob<sup>8</sup>, P. Andre<sup>1,2,3,11</sup>, V. Lotteau<sup>1,2,3,11</sup>, C. Rabourdin-Combe<sup>1,2,3,\*</sup> and P. O. Vidalain<sup>4,\*</sup>

<sup>1</sup>INSERM U851, Lyon, <sup>2</sup>IFR128-BioSciences, Université Lyon 1, <sup>3</sup>Université de Lyon, Lyon, <sup>4</sup>Laboratoire de Génomique Virale et Vaccination, Institut Pasteur, CNRS URA3015, Paris, <sup>5</sup>INRA UMR754, Lyon, <sup>6</sup>Modul-Bio, Marseille, <sup>7</sup>CNRS, UMR5558, Laboratoire de Biologie et de Biométrie Évolutive, Lyon, <sup>8</sup>Unité de Génétique, Papillomavirus et Cancer Humain, Institut Pasteur, Paris, <sup>9</sup>RDP, UMR 5667 INRA/CNRS/ENSL/UCB-Lyon I, <sup>10</sup>IBCP, CNRS UMR5086 and <sup>11</sup>Hospices Civils de Lyon, Hôpital de la Croix-Rousse, Laboratoire de virologie, Lyon, France

Received August 5, 2009; Revised September 18, 2009; Accepted October 18, 2009

## ABSTRACT

Large collections of protein-encoding open reading frames (ORFs) established in a versatile recombination-based cloning system have been instrumental to study protein functions in high-throughput assays. Such 'ORFeome' resources have been developed for several organisms but in virology, plasmid collections covering a significant fraction of the virosphere are still needed. In this perspective, we present ViralORFeome 1.0 (<http://www.viralorfeome.com>), an open-access database and management system that provides an integrated set of bioinformatic tools to clone viral ORFs in the Gateway<sup>®</sup> system. ViralORFeome provides a convenient interface to navigate through virus genome sequences, to design ORF-specific cloning primers, to validate the sequence of generated constructs and to browse established collections of virus ORFs. Most importantly, ViralORFeome has been designed to manage all possible variants or mutants of a given ORF so that the cloning procedure can be applied to any emerging virus strain. A subset of plasmid constructs generated with ViralORFeome platform has

been tested with success for heterologous protein expression in different expression systems at proteome scale. ViralORFeome should provide our community with a framework to establish a large collection of virus ORF clones, an instrumental resource to determine functions, activities and binding partners of viral proteins.

## INTRODUCTION

The number of viral genomic sequences available in public databases has increased exponentially, opening new perspectives to understand genetic basis and functional mechanisms that underlie virus replication, pathogenesis and evolution. In particular, this enabled to establish for each virus a list of potential regulatory and expressed sequences, a framework often referred as the 'parts list' of biological systems (1). Current investigations aim at understanding how these viral components act upon each other and interact with host macromolecules to carry on viral replication and spreading. To reach such a system view of virus cycles, more functional analyses of viral components are necessary, especially in the field of virus–host molecular interactions (2,3). To address this question, a large collection of viral open reading frames (ORFs) established in a recombination-based cloning

\*To whom correspondence should be addressed. Tel: +33 4 37 28 23 42; Fax: +33 4 37 28 23 21; Email: [chantal.rabourdin@inserm.fr](mailto:chantal.rabourdin@inserm.fr)  
Correspondence may also be addressed to P.O. Vidalain. Tel: +33 1 45 68 87 73; Fax: +33 1 40 61 31 67; Email: [vidalain@pasteur.fr](mailto:vidalain@pasteur.fr)  
Present address:

V. Navratil, Université de Lyon, CNRS/ENSL/UCB-Lyon I, Centre de RMN à Très Hauts Champs, Villeurbanne, France.

The authors wish it to be known that, in their opinion, the first three authors and the last two authors should be regarded as joint First Authors.

© The Author(s) 2009. Published by Oxford University Press.

This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/2.5/uk/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

system allowing their mass transfer into various functional assays would be extremely helpful. Such ORF collections, often referred as 'ORFeomes', have been developed for human and few other organisms and represent great resources to explore protein functions in a large-scale setting (4–9). Recombination-based cloning technologies like the Gateway® system enable the mass cloning of polymerase chain reaction (PCR)-amplified ORFs into a 'donor' vector to create 'entry' clones. Once entry clones have been established, ORFs can be easily recombined into different 'destination' vectors that allow protein expression. So far, only few viral ORFeomes have been built and they are dedicated to a single virus, arguing that much more needs to be achieved (10–13).

Because of their small-sized genomes, most viruses exhibit a limited number of ORFs. Therefore, building a viral ORFeome collection covering a significant number of viral pathogens can appear like a manageable project. However, this eventually becomes a daunting task when considering virus strains and polymorphisms that affect viral proteins (Table 1). Nonetheless, these variations cannot be neglected since they often determine virulence and/or host adaptation. For example, a unique amino acid mutation in the polymerase of poliovirus can alter its processivity, turning a highly pathogenic virus into an attenuated strain that can eventually be used as a vaccine (14). Similarly, a Semliki Forest virus strain encoding an nsP2 protein, with a single amino acid mutation in its nuclear localization signal, is impaired for the control of type I interferon response and strongly attenuated *in vivo* (15). Thus, specific care must be taken when cloning a virus ORFeome, since an accurate identification of the strain that is used as a template is absolutely critical, both in terms of genotype and phenotype. For this reason, wild-type virus strains corresponding to primary isolates will be generally preferred to culture-adapted laboratory strains that often exhibit an altered phenotype *in vivo*. Such biological samples are often difficult to obtain, since it requires access to patients, medical-care facilities and laboratories of an appropriate biosafety level. With such constraints to obtain suitable viral RNA or DNA templates, a collaborative effort between virology laboratories is needed to build a comprehensive viral ORFeome resource. This motivated the development of ViralORFeome 1.0, an open-access database and management system designed to assist academic laboratories in the development of a viral ORFeome collection using a recombination-based cloning technology.

## DESIGN AND IMPLEMENTATION

ViralORFeome is a database for creating and managing viral ORF clones. It includes a web interface to search

and display viral sequences and annotations downloaded from GenBank. When users have identified the viral ORF sequences that they want to clone, clicking on target sequences automatically creates virtual clones in ViralORFeome database together with suitable cloning primers. These primers are designed for PCR amplification and cloning of viral ORFs using the Gateway® system. To determine if 'entry' clones produced at the bench exhibit expected sequences, sequence traces can be uploaded in ViralORFeome, automatically aligned, and compared to corresponding virtual clones. Finally, ViralORFeome database allows users to manage their collection of viral ORF clones, keep track of expression plasmids derived from 'entry' clones and share plasmids with other laboratories. System architecture and data flow are depicted in Figure 1.

### ViralORFeome architecture

The World Wide Web server for ViralORFeome is Apache (<http://www.apache.org/>) with Hypertext Pre-processor (PHP, <http://www.php.net>) and Asynchronous JavaScript and XML (AJAX) technologies. The relational database management system for ViralORFeome is PostgreSQL (<http://www.postgresql.org/>). ViralORFeome Entity-Relationship (ER) model was divided into five interconnected parts: 'Taxonomy', 'Genome', 'ORFeome', 'Interactome' and 'Users' schemes (Supplementary Figure S1).

### ViralORFeome data source

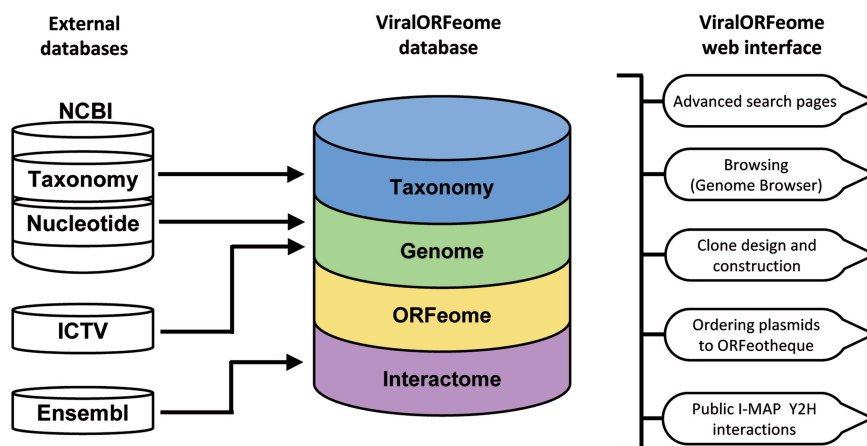
ViralORFeome 1.0 is based on conventional sequence and virus databases. Virus and host organism classification was retrieved from the taxonomy database at National Center for Biotechnology Information [NCBI; <ftp://ftp.ncbi.nih.gov/pub/taxonomy/>; (16)]. Complete viral CDS and associated annotations were downloaded from the CoreNucleotide division of GenBank [<http://www.ncbi.nlm.nih.gov/sites/entrez?db=nucleotide>; (17)] using Entrez Programming Utilities (eUtils) at NCBI (see Supplementary Data). ICTV names were obtained from the most recent publication of the International Committee on Taxonomy of Viruses using Perl scripts (18). Finally, host gene and protein sequence annotations were extracted from Ensembl database (19).

### ViralORFeome interface

ViralORFeome database can be accessed at: <http://www.viralorfeome.com>. When selecting 'Virus' option in the menu bar, ViralORFeome web interface allows users to navigate throughout viral sequences retrieved from GenBank by entering criteria such as virus species, taxon IDs or accession numbers (Supplementary Figure S2). ViralORFeome can also execute a BLASTn or BLASTp search among available viral sequences (20). Virus genomes and coding sequences of interest are visualized with an integrated genome browser adapted from GBrowse that provides a synthetic view of genomic sequence features (21). When users have identified a viral ORF sequence that they want to clone, clicking on this sequence automatically generates cloning primers and

**Table 1.** Viral sequences available in ViralORFeome 1.0 (GenBank release 09-2009)

Viral species	Virus strains	Complete CDS	Annotated proteins
2290	89 839	320 657	377 967



**Figure 1.** Overview of ViralORFeome database architecture. Data from external public databases (left panel) have been integrated in ViralORFeome database (middle panel): genomic and taxonomic data have been extracted from NCBI and ICTV using Perl scripts. Host protein sequences and annotations have been obtained from Ensembl. ViralORFeome web interface (right panel) enables the design of ORF clones using viral CDS (Coding DNA Sequences) available in GenBank as templates. Corresponding plasmids are stored in ORFeothèque (i.e. viral ORFs physical library), and can be requested using the same interface. Raw data relative to Y2H experiments performed with viral ORF clones have been analyzed with an IST pipeline [pISTil, (31)] and are stored in ViralORFeome.

virtual clones in ViralORFeome database (Figure 2). ORF-specific Gateway® primers are designed using the Oligonucleotide Selection Program (OSP) software adapted for recombinational cloning (22). By default, primers are designed to clone full-length ORFs from ATG to STOP codon according to GenBank annotations. However, users can specify 5'- and 3'-coordinates to clone ORF fragments corresponding to specific domains or even upload manually designed primers. Once primers have been validated, virtual clones are generated and used as entry points in ViralORFeome to store and access all information relative to the viral ORF constructs (e.g. primers, sequencing traces, potential mutations, and comments).

Importantly, when users want to clone some variant of an ORF that is not defined in GenBank, and therefore is not available in ViralORFeome, an interface allows for the creation of a 'variant' clone that is anchored to the most similar annotated sequence found by BLAST in the database (Supplementary Figure S3). This ensures a maximum of flexibility when viral sequences are known and characterized but not readily available in GenBank, a frequent situation when working with primary isolates. This also facilitates the management of mutants generated on purpose to study the role of specific residues in a protein's function.

In this recombination-based cloning pipeline, viral ORFs amplified by PCR or RT-PCR are cloned by *in vitro* recombination into donor vectors such as pDONR207 or pDONR223 (see Supplementary Data for detailed protocols). After transformation, a construct can be purified either from a mini-pool of bacteria colonies to keep the sequence diversity of the original template (e.g. when working with viral quasi-species) or from a single isolated colony. These two cloning strategies, referred as 1.0 and 2.0, respectively (5), are manageable using ViralORFeome interface (Figure 2).

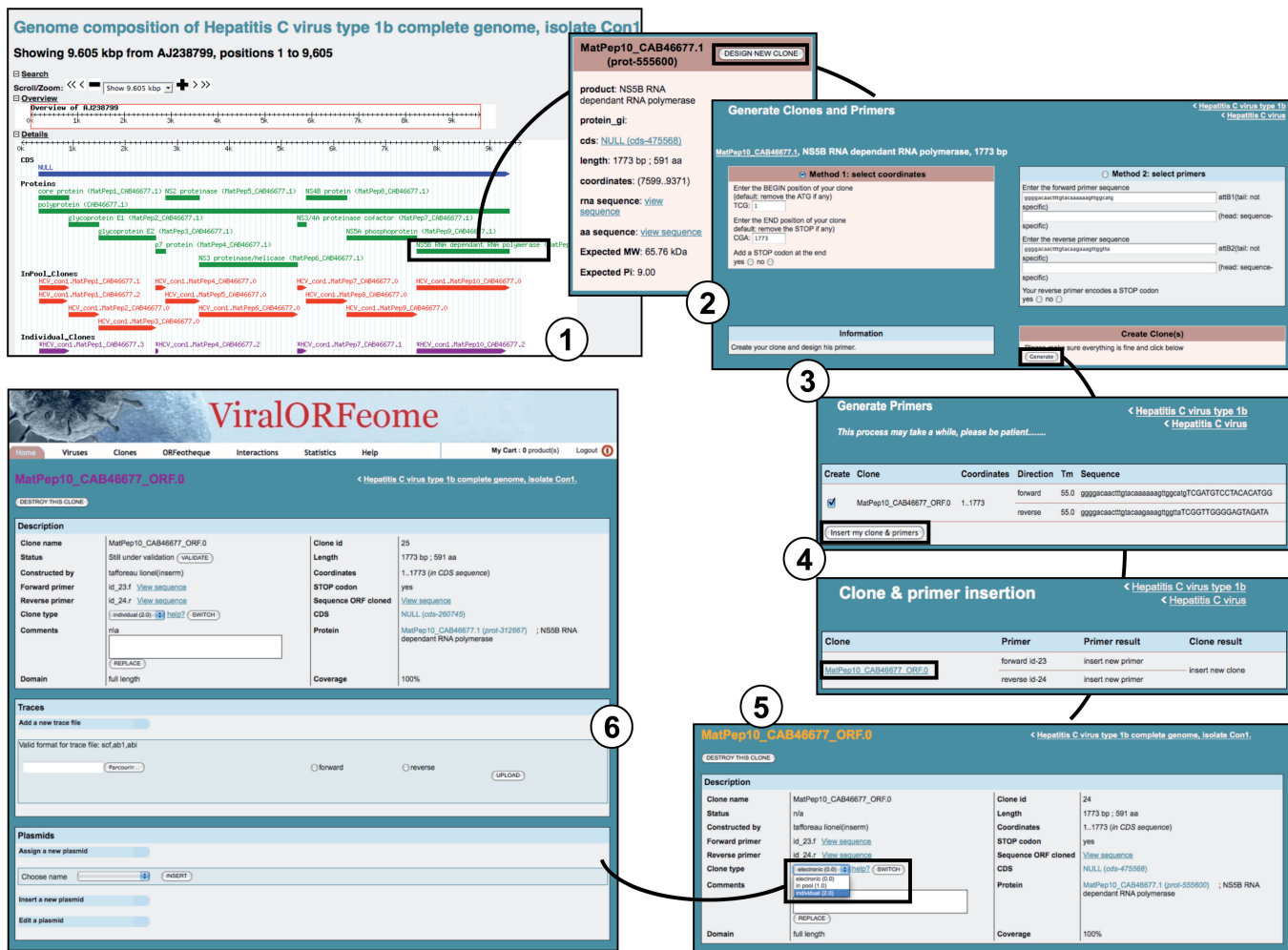
To validate 'entry' constructs by sequencing, ViralORFeome automatically aligns uploaded sequence traces onto virtual clones using either an extended Libalign Perl tool or Phred and T-Coffee programs that build first a contig of sequence traces before the alignment is performed (23,24). Validated 'entry' constructs can be subsequently used to recombine viral ORFs into various Gateway®-compatible destination vectors to achieve protein expression. ViralORFeome allows users to keep track of the different expression constructs that have been established for each viral ORF.

Users can browse the set of viral ORF clones already established in ViralORFeome by selecting the 'clones' option in the menu bar. Queries can be performed by entering criteria such as virus species, GenBank protein accession numbers or ViralORFeome clone IDs. Viral ORF clones can also be searched by nucleotide or protein BLAST. Until now, three laboratories involved in the Infection-MAPPING project (I-MAP) have used ViralORFeome interface and generated a collection of 528 viral ORFs cloned into pDONR vectors. Among them, 145 have already been stored in a dedicated repository for viral ORFeome resource (ORFeothèque; Hospices Civils de Lyon), and 134 are available upon request under material transfer agreement via ViralORFeome. This set of viral ORF clones can be searched selecting the 'ORFeothèque' option in the menu bar.

## EXPERIMENTAL VALIDATION

To validate the viral ORF clone collection built using ViralORFeome pipeline, we tested a subset of 66 ORFs (Figure 3a) from three *Alphaviruses* (Chikungunya, Sindbis and Semliki Forest viruses), one *Flavivirus* (yellow fever virus), one *Hepadnavirus* (hepatitis B virus) and four *Paramyxoviruses* (measles virus wild-type and vaccine strains, mumps, Tioman and Nipah viruses).



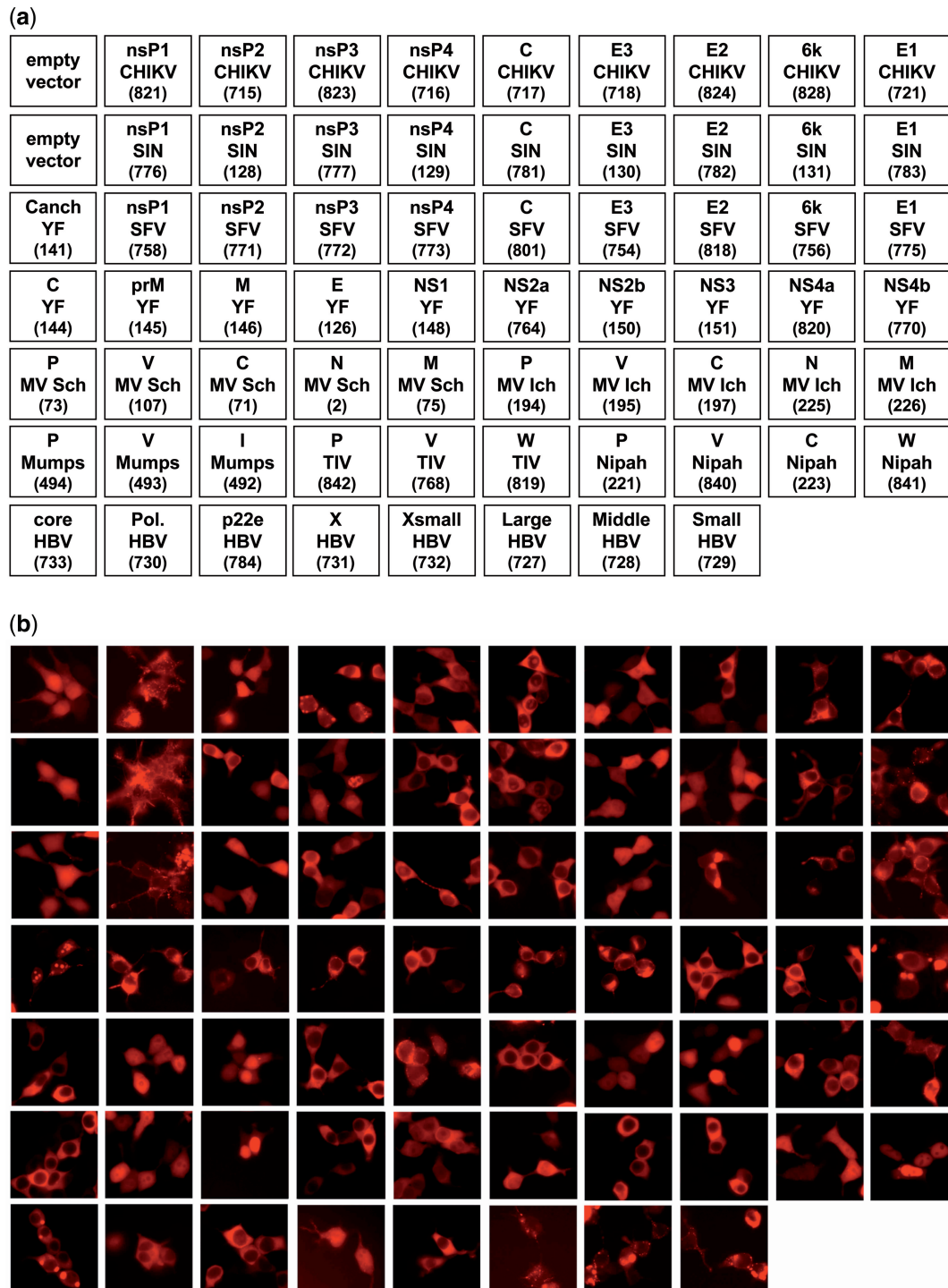


**Figure 2.** Building a viral ORF collection using ViralORFeome interface. Viral sequences and annotations from GenBank are visualized with a genome browser that provides a synthetic view of sequence features (1). CDS are shown in blue and proteins in green. Users can design a new clone by clicking on a viral protein of interest (2). By default, ViralORFeome will anchor cloning primers at the extremities of selected ORFs (Method 1), but user can specify 5'- and 3'-coordinates and clone ORF fragments corresponding to specific domains. Users can also upload manually designed primers (Method 2). ViralORFeome will automatically design Gateway<sup>®</sup> cloning primers (3) and after validation (4), a virtual clone is created in the database (5). Users need to select between two cloning strategies, 1.0 ('in pool') or 2.0 ('individual clone'), before they can access a webpage where all information relative to the construct are stored (6). This includes clone coordinates, primers, sequence and comments (upper panel), sequencing traces and alignments (middle panel), and available entry and destination vectors to achieve viral ORF expression (lower panel). When back to the genome browser (1), viral ORF clones are displayed in red (1.0 constructs) or purple (2.0 constructs).

ORFs were expressed in fusion downstream of the red fluorescent protein Cherry and constructs transfected in human cells to achieve subcellular localization (Figure 3b). We observed a specific localization pattern distinct from Cherry alone for most constructs, demonstrating that viral ORFs are properly expressed in this system. In addition, and even if the N-terminal tag was expected to alter the localization of proteins like viral envelopes, observed localization patterns were generally consistent with literature. For example, expression of nsP1 not only from Semliki Forest virus and Sindbis virus (25) but also Chikungunya virus induced characteristic filopodia-like extensions in transfected cells. The capsid of yellow fever virus (C) was found in the nucleus where it accumulated in a dot-like pattern, reminiscent of

nucleoli, as reported for C of dengue and Japanese encephalitis viruses (26,27).

As numerous viral proteins are known to inhibit the host immune response, the 66 ORFs were also tested for their ability to block signaling downstream of two key antiviral cytokines: Interferon- $\beta$  (IFN- $\beta$ ) and tumor necrosis factor- $\alpha$  (TNF- $\alpha$ ). Each ORF was expressed in fusion downstream of the 3xFLAG tag and, using luciferase reporter constructs, tested for its ability to inhibit IFN- $\beta$  or TNF- $\alpha$  signaling (Figure 3c and d). Both pathways were inhibited by nsP2 from Chikungunya and Semliki forest viruses. This is consistent with previous reports that used mutant viruses or replicons to demonstrate that nsP2 from Old World *Alphaviruses* induces a transcriptional shutoff and



**Figure 3.** Expression and functional validation of plasmid constructs generated with ViralORFeome database. (a) Matrix map of virus ORF constructs that have been tested in expression and functional assays. Clone IDs in ViralORFeome database are indicated between parentheses. CHIKV, Chikungunya virus; SIN, Sindbis virus; SFV, Semliki Forest virus; YF, Yellow Fever virus; MV Sch, Schwarz vaccine strain of measles virus; MV Ich, Ichinose wild-type strain of measles virus; TIV, Tioman virus; HBV, hepatitis B virus. (b) Virus ORFs were recombined from entry vectors pDONR207 or pDONR223 in a Gateway<sup>®</sup>-compatible expression vector to be expressed in fusion with the red-fluorescent protein Cherry. Plasmids were transfected in HEK-293T cells and subcellular localization determined 24 h later. (c and d) The same virus ORFs were recombined in an expression vector to be expressed in fusion with the 3×FLAG tag. Constructs were co-transfected in HEK-293T cells together with a reporter plasmid encoding luciferase downstream of a promoter containing either IFN- $\alpha/\beta$  (pISRE-Luc) or NF- $\kappa$ B (pNF- $\kappa$ B-Luc) response elements. A CMV-Renilla plasmid was also co-transfected and used as an internal control for transfection efficiency and cell viability. After transfection, cells were incubated for 24 h in the presence of IFN- $\beta$  (c) or TNF- $\alpha$  (d) to activate ISRE or NF- $\kappa$ B response elements, respectively. Cells in position A1 and B1 correspond to negative and positive controls that were respectively left untreated or stimulated with IFN- $\beta$  or TNF- $\alpha$ . Relative luciferase activity was determined using a chemiluminescent substrate, and results expressed in relative percentage to positive control. Data show one representative experiment out of two.

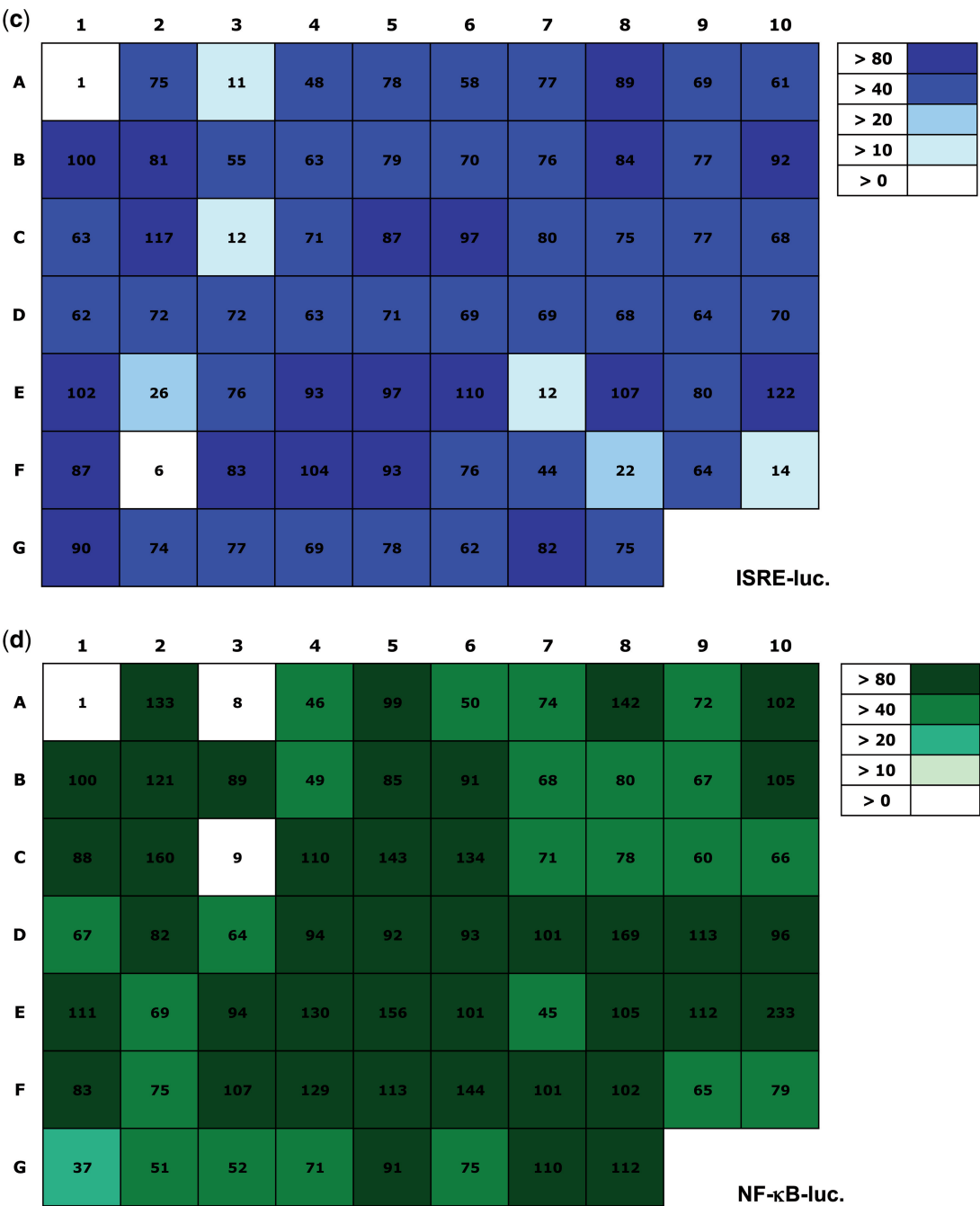


Figure 3. Continued.

controls the host antiviral response (15,28). In contrast, nsP2 derived from a Sindbis virus infectious cDNA clone (29) did not localize in the nucleus and failed to block signaling. This supports previous reports showing that nsP2 from *Alphaviruses* must be nuclear to control the antiviral response (15). We also confirmed that V proteins of measles and mumps viruses and V and W proteins of Nipah virus block IFN-β signaling (30). Interestingly, the V protein of Tioman virus was unable to do so, suggesting that this virus infecting flying foxes (*Pteropus* genus) is not adapted to human cells. Altogether, observed

localization patterns and functional assays validate the clone collection that was generated with ViralORFeome platform. Furthermore, these results illustrate how large collections of viral ORF clones established in a versatile cloning system provide access to reverse proteomic platforms and large-scale functional assays.

SUMMARY AND FUTURE DEVELOPMENTS

In conclusion, ViralORFeome is the first open-access database that provides an integrated set of bioinformatic

tools to build a collection of viral ORFs clones in a versatile system suitable for reverse proteomic experiments. In this perspective, ViralORFeome was especially designed to handle the diversity of virus strains, variants and species. As shown here, our cloning pipeline has been validated using functional assays and a collection of 528 viral ORFs has been generated in the Gateway® system. As collaborative efforts between virology laboratories are required to establish an ORFeome collection covering most viral genera and species, we believe ViralORFeome will provide the community with a framework to sustain this global effort.

In the near future, our objective will be to motivate more laboratories to join this program. In addition, we would like to implement new modules allowing users to store and visualize all kind of functional data generated with viral ORFs from the collection. This virus clone collection was primarily developed to map virus–host protein–protein interactions using the yeast two-hybrid (Y2H) system as a collaborative effort between co-authors of this manuscript. Current storage and display of virus–host interaction data obtained by Y2H screening of HCV proteins constitute a first attempt to reach this goal. The ‘Interaction’ option of the menu bar allows users to select HCV ORFs and access all interacting cellular cDNA clones identified by Y2H. Whereas filtered data sets have already been published (12), ViralORFeome interface allows users to access Y2H raw data to filter these results according to their own quality criteria. Other functional data such as subcellular localization or interference with immune response pathways should be implemented.

## ACCESS TO THE DATABASE

Public access to the ViralORFeome is available at: <http://www.viralorfeome.com>. Registration is not necessary, and use of the database is free. Users who want to generate viral ORF clones that are tagged with their institutional name can create an account to register.

## SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

## ACKNOWLEDGEMENTS

We thank all members of PF1-Pasteur Genopole sequencing core facility, in particular C. Bouchier and C. Gouyette. We thank Drs Marie-Louise Michel, Charles M. Rice, Kaoru Takeuchi, T. Fabian Wild and Ali Amara for providing RNA or DNA templates used to build viral ORF clones. We also thank Eric Coissac for providing SQL codes.

## FUNDING

The Institut National de la Santé et de la Recherche Médicale; the Institut National de la Recherche Agronomique; the Association de la Recherche contre le

Cancer (3731XA0531F and 4867); the Ligue Nationale Contre le Cancer (RS07/75-75); the Agence Nationale de la Recherche (EPI-HPV-3D); the French Ministry of Industry; the Institut Pasteur; the Centre National de la Recherche Scientifique (Maladies Infectieuses Emergentes to P.O.V., G.C. and F.T.). Funding for open access charge: Institut Pasteur.

*Conflict of interest statement.* None declared.

## REFERENCES

- Brasch, M.A., Hartley, J.L. and Vidal, M. (2004) ORFeome cloning and systems biology: standardized mass production of the parts from the parts-list. *Genome Res.*, **14**, 2001–2009.
- Uetz, P., Rajagopala, S.V., Dong, Y.A. and Haas, J. (2004) From ORFeomes to protein interaction maps in viruses. *Genome Res.*, **14**, 2029–2033.
- Navratil, V., de Chassey, B., Meyniel, L., Delmotte, S., Gautier, C., Andre, P., Lotteau, V. and Rabourdin-Combe, C. (2009) VirHostNet: a knowledge base for the management and the analysis of proteome-wide virus–host interaction networks. *Nucleic Acids Res.*, **37**, D661–D668.
- Reboul, J., Vaglio, P., Rual, J.F., Lamesch, P., Martinez, M., Armstrong, C.M., Li, S., Jacotot, L., Bertin, N., Janky, R. *et al.* (2003) C. elegans ORFeome version 1.1: experimental verification of the genome annotation and resource for proteome-scale protein expression. *Nat. Genet.*, **34**, 35–41.
- Rual, J.F., Hirozane-Kishikawa, T., Hao, T., Bertin, N., Li, S., Dricot, A., Li, N., Rosenberg, J., Lamesch, P., Vidalain, P.O. *et al.* (2004) Human ORFeome version 1.1: a platform for reverse proteomics. *Genome Res.*, **14**, 2128–2135.
- Dricot, A., Rual, J.F., Lamesch, P., Bertin, N., Dupuy, D., Hao, T., Lambert, C., Hallez, R., Delroisse, J.M., Vandenhoute, J. *et al.* (2004) Generation of the Brucella melitensis ORFeome version 1.1. *Genome Res.*, **14**, 2201–2206.
- Schroeder, B.K., House, B.L., Mortimer, M.W., Yurgel, S.N., Maloney, S.C., Ward, K.L. and Kahn, M.L. (2005) Development of a functional genomics platform for Sinorhizobium meliloti: construction of an ORFeome. *Appl. Environ. Microbiol.*, **71**, 5858–5864.
- Matsuyama, A., Arai, R., Yashiroda, Y., Shirai, A., Kamata, A., Sekido, S., Kobayashi, Y., Hashimoto, A., Hamamoto, M., Hiraoka, Y. *et al.* (2006) ORFeome cloning and global analysis of protein localization in the fission yeast Schizosaccharomyces pombe. *Nat. Biotechnol.*, **24**, 841–847.
- Brandner, C.J., Maier, R.H., Henderson, D.S., Hintner, H., Bauer, J.W. and Onder, K. (2008) The ORFeome of Staphylococcus aureus v 1.1. *BMC Genomics*, **9**, 321.
- Calderwood, M.A., Venkatesan, K., Xing, L., Chase, M.R., Vazquez, A., Holthaus, A.M., Ewence, A.E., Li, N., Hirozane-Kishikawa, T., Hill, D.E. *et al.* (2007) Epstein–Barr virus and virus human protein interaction maps. *Proc. Natl Acad. Sci. USA*, **104**, 7606–7611.
- von Brunn, A., Teepe, C., Simpson, J.C., Pepperkok, R., Friedel, C.C., Zimmer, R., Roberts, R., Baric, R. and Haas, J. (2007) Analysis of intraviral protein–protein interactions of the SARS coronavirus ORFeome. *PLoS ONE*, **2**, e459.
- de Chassey, B., Navratil, V., Tafforeau, L., Hiet, M.S., Aublin-Gex, A., Agaoglu, S., Meiffren, G., Pradezynski, F., Faria, B.F., Chantier, T. *et al.* (2008) Hepatitis C virus infection protein network. *Mol. Syst. Biol.*, **4**, 230.
- Zhang, L., Villa, N.Y., Rahman, M.M., Smallwood, S., Shattuck, D., Neff, C., Dufford, M., Lanchbury, J.S., Labaer, J. and McFadden, G. (2009) Analysis of vaccinia virus–host protein–protein interactions: validations of yeast two-hybrid screenings. *J. Proteome Res.*, **8**, 4311–4318.
- Vignuzzi, M., Wendt, E. and Andino, R. (2008) Engineering attenuated virus vaccines by controlling replication fidelity. *Nat. Med.*, **14**, 154–161.



15. Breakwell, L., Dosenovic, P., Karlsson Hedestam, G.B., D'Amato, M., Liljestrom, P., Fazakerley, J. and McInerney, G.M. (2007) Semliki Forest virus nonstructural protein 2 is involved in suppression of the type I interferon response. *J. Virol.*, **81**, 8677–8684.
16. Jenuth, J.P. (2000) The NCBI. Publicly available tools and resources on the Web. *Methods Mol. Biol.*, **132**, 301–312.
17. Sayers, E.W., Barrett, T., Benson, D.A., Bryant, S.H., Canese, K., Chetvernin, V., Church, D.M., DiCuccio, M., Edgar, R., Federhen, S. *et al.* (2009) Database resources of the National Center for Biotechnology Information. *Nucleic Acids Res.*, **37**, D5–D15.
18. Buchen-Osmond, C. (1997) Further progress in ICTVdB, a universal virus database. *Arch. Virol.*, **142**, 1734–1739.
19. Hubbard, T.J., Aken, B.L., Ayling, S., Ballester, B., Beal, K., Bragin, E., Brent, S., Chen, Y., Clapham, P., Clarke, L. *et al.* (2009) Ensembl 2009. *Nucleic Acids Res.*, **37**, D690–D697.
20. Altschul, S.F., Madden, T.L., Schaffer, A.A., Zhang, J., Zhang, Z., Miller, W. and Lipman, D.J. (1997) Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. *Nucleic Acids Res.*, **25**, 3389–3402.
21. Stein, L.D., Mungall, C., Shu, S., Caudy, M., Mangone, M., Day, A., Nickerson, E., Stajich, J.E., Harris, T.W., Arva, A. *et al.* (2002) The generic genome browser: a building block for a model organism system database. *Genome Res.*, **12**, 1599–1610.
22. Hillier, L. and Green, P. (1991) OSP: a computer program for choosing PCR and DNA sequencing primers. *PCR Methods Appl.*, **1**, 124–128.
23. Ewing, B., Hillier, L., Wendl, M.C. and Green, P. (1998) Base-calling of automated sequencer traces using phred. I. Accuracy assessment. *Genome Res.*, **8**, 175–185.
24. Notredame, C., Higgins, D.G. and Heringa, J. (2000) T-Coffee: a novel method for fast and accurate multiple sequence alignment. *J. Mol. Biol.*, **302**, 205–217.
25. Laakkonen, P., Auvinen, P., Kujala, P. and Kaariainen, L. (1998) Alphavirus replicase protein NSP1 induces filopodia and rearrangement of actin filaments. *J. Virol.*, **72**, 10265–10269.
26. Wang, S.H., Syu, W.J., Huang, K.J., Lei, H.Y., Yao, C.W., King, C.C. and Hu, S.T. (2002) Intracellular localization and determination of a nuclear localization signal of the core protein of dengue virus. *J. Gen. Virol.*, **83**, 3093–3102.
27. Mori, Y., Okabayashi, T., Yamashita, T., Zhao, Z., Wakita, T., Yasui, K., Hasebe, F., Tadano, M., Konishi, E., Moriishi, K. *et al.* (2005) Nuclear localization of Japanese encephalitis virus core protein enhances viral replication. *J. Virol.*, **79**, 3448–3458.
28. Garmashova, N., Gorchakov, R., Volkova, E., Paessler, S., Frolova, E. and Frolov, I. (2007) The Old World and New World alphaviruses use different virus-specific proteins for induction of transcriptional shutoff. *J. Virol.*, **81**, 2472–2484.
29. Rice, C.M., Levis, R., Strauss, J.H. and Huang, H.V. (1987) Production of infectious RNA transcripts from Sindbis virus cDNA clones: mapping of lethal mutations, rescue of a temperature-sensitive marker, and in vitro mutagenesis to generate defined mutants. *J. Virol.*, **61**, 3809–3819.
30. Fontana, J.M., Bankamp, B. and Rota, P.A. (2008) Inhibition of interferon induction and signaling by paramyxoviruses. *Immunol. Rev.*, **225**, 46–67.
31. Pellet, J., Meyniel, L., Vidalain, P.O., de Chassey, B., Tafforeau, L., Lotteau, V., Rabourdin-Combe, C. and Navratil, V. (2009) pISTil: a pipeline for yeast two-hybrid Interaction Sequence Tags identification and analysis. *BMC Res. Notes*, **2**, 220.